

ProyectoMachine Learning I

¡Hola a todos!

En este proyecto quise enfocarme en hacer una pequeña simulación de lo que pasa realmente en un proyecto de ML, por ello recurrí a buscar datos reales, en formato csv. Tiene muchos targets (Posibles X), no deben usar todas, sino seleccionen la que les interese. Deben entregar el ipynb con su modelo, el pdf de su ipynb y un pdf explicando lo que hicieron entre otros requerimientos que encontrarán especificados más abajo.

Aplicar un modelo de clasificación a los datos de Covid-19 dando un pronóstico de recuperación / muerte o leve / grave (scratch y sklearn)

Los datos se encuentran en el archivo "COVID19_open_line_list.csv", que como pueden imaginarse por el título contiene información de pacientes de la reciente pandemia; estos datos no están completamente limpios, ni listos para que implementen su modelo, les entregó el csv directo de la fuente. NO todas las columnas deben ser usadas, ya que ciertos datos son inútiles para poder predecir la resolución.

Deben entregarme el ipynb y pdf del mismo con su modelo y una predicción hecha por ustedes. Independientemente de comentar el código deben ir explicando paso a paso el proceso, ej., Dividir en test set y training test seguido por una celda con esa parte del código. Además deberán entregar un escrito con formato PDF donde den una introducción de los datos, explicar el algoritmo de clasificación de su elección y justificar las decisiones tomadas.

Como ayuda les adjunto una libreta de jupyter titulada "Proyecto_pt1_coronavirus.ipynb" donde realice la carga de los datos y los primeros pasos del manejo de los datos.

Rúbrica:

	Puntos
Implementación scratch	3
Implementación sklearn	2
Predicción del modelo	3
PDF explicando su proyecto	2
Comparación scratch/sklearn	1
Código comentado	+1
Total	11

Cualquier duda o pregunta pueden comunicarse conmigo.

Saludos, Adriana.